

# (12) UK Patent Application (19) GB (11) 2 365 262 (13) A

(43) Date of A Publication 13.02.2002

(21) Application No 0018027.3

(22) Date of Filing 21.07.2000

(71) Applicant(s)  
Telefonaktiebolaget L M Ericsson (publ)  
(Incorporated in Sweden)  
SE-126 25 Stockholm, Sweden

(72) Inventor(s)  
Sten Minör  
Magnus Tillgren  
Patrik Olsson

(74) Agent and/or Address for Service  
Hasettine Lake & Co  
Imperial House, 15-19 Kingsway, LONDON,  
WC2B 6UD, United Kingdom

(51) INT CL<sup>7</sup>

H04L 12/64 // H04L 12/66, H04M 1/27 3/493 7/00  
11/00

(52) UK CL (Edition T )  
H4L LRAB

(56) Documents Cited  
EP 1041779 A2 JP 010084425 A

(58) Field of Search  
UK CL (Edition S ) H4L LRAB LRAD LRAX  
INT CL<sup>7</sup> G01L 5/06, G06F 3/16 11/00 13/00 15/163,  
G10L 15/26, H04L 12/02 12/18 12/28 12/46 12/54  
12/56 12/58 12/64 12/66, H04M 1/00 1/26 3/00 3/42  
3/493 11/00 11/06 15/00 15/16, H04Q 3/00 11/04  
Online: WPI EPODOC JAPIO

(54) Abstract Title

**A network location address is made up of address components which are associated with reference voice signals which match voice signals received by a mobile**

(57) A method of accessing a network using a network access application is disclosed. The network has a plurality of locations which have respective location addresses in the network. The method includes defining a plurality of address components, combinations of which forming respective location addresses, assigning reference voice signals to the address components, comparing input voice signals with the reference voice signals and producing a network access address location on the basis of address components associated with matched input and reference voice signals.

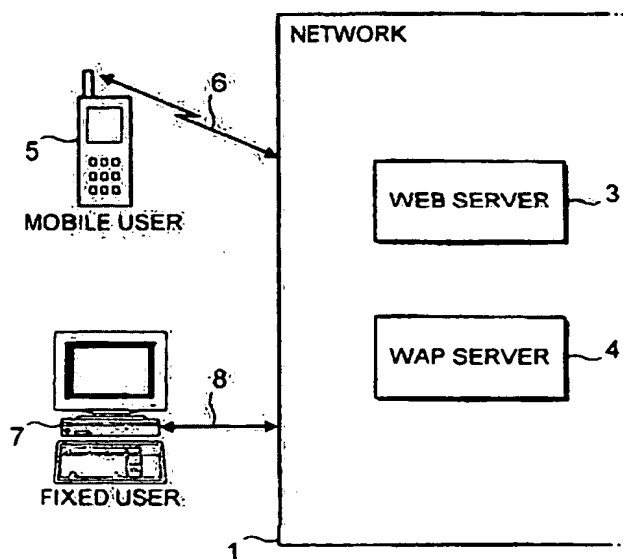


FIG. 1

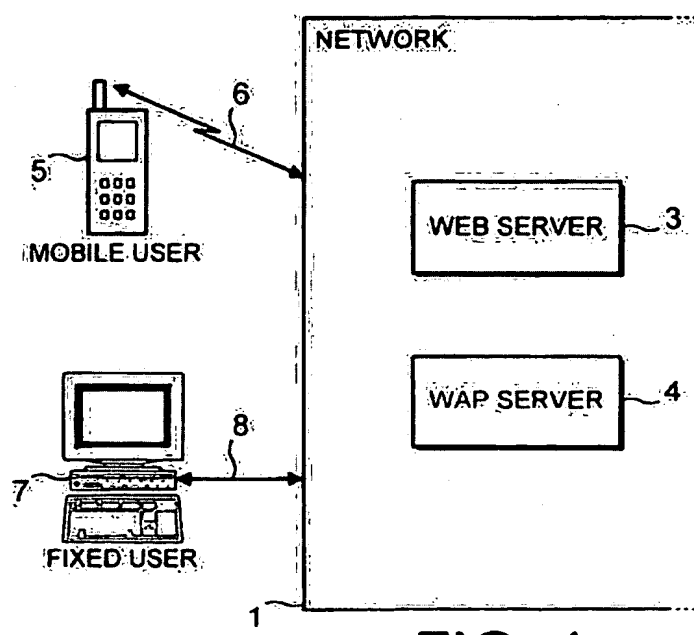


FIG. 1

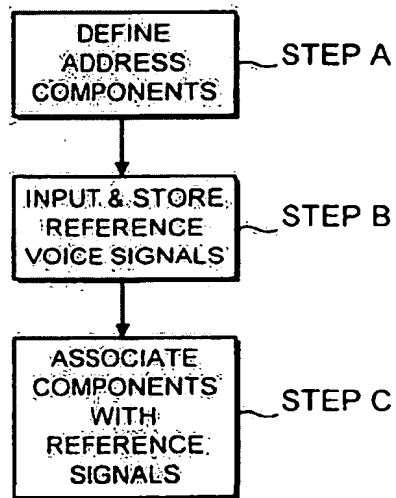
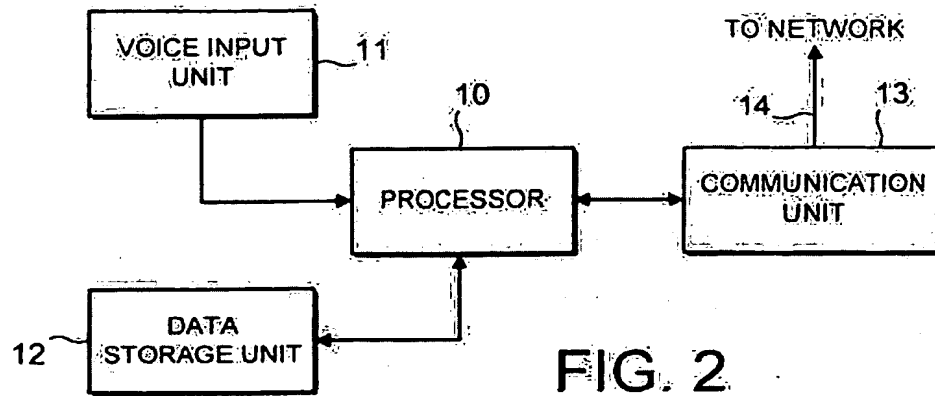


FIG. 3

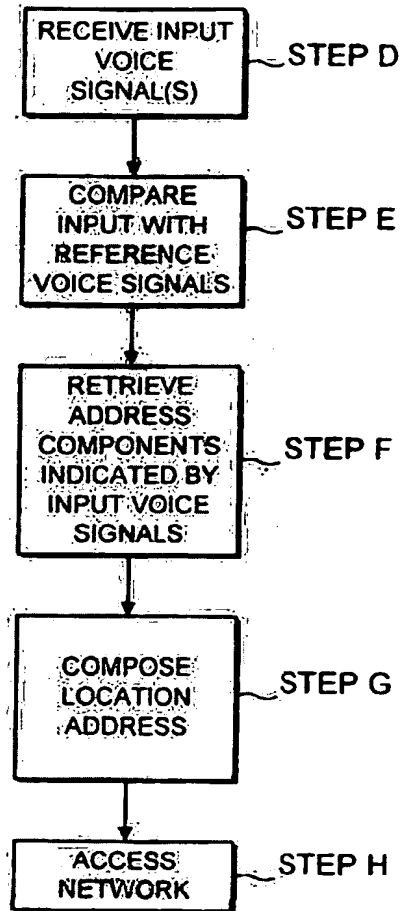


FIG. 4

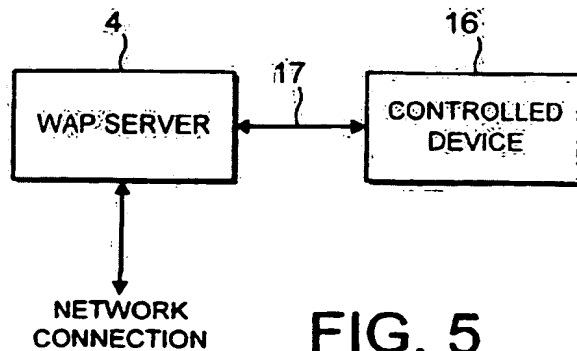


FIG. 5

COMMUNICATION SYSTEMS

The present invention relates to communication systems, and in particular to methods of communicating with a network.

Background of the present Invention

Access to data networks, for example the internet, is typically achieved using a conventional personal computer (PC) terminal, which has a large high resolution display and relatively high data transmission bandwidth. In order to communicate with the network concerned the user of the PC typically inputs a network address via a keyboard or other input device, such as a mouse. These devices enable long and complex network addresses to be input to the PC relatively easily so that network locations may be accessed.

A recent development in network, and particularly the internet, access is the standardization of a wireless application protocol (WAP), using the so-called "wireless markup language (WML)". The WAP system enables terminals with small displays, limited processing power and low data transmission bandwidth (for example digital mobile telephones and similar devices) to access and control services and content in a network such as the internet. WAP is a layered communications protocol that includes network layers as well as an application environment including a browser (network communication software). WML uses simple syntax and limited vocabulary in order to make it suitable for controlling communications with a network by a user having access to a portable device.

Whilst the ability to use smaller terminals can be a major convenience to a user, it can be inconvenient to access complex network addresses using a relatively simple keypad input device. For example, a mobile  
5 telephone has a keypad primarily intended for inputting telephone numbers, and so input of other characters can become complex and tiresome. It has been proposed to overcome this disadvantage by providing voice control of the user interface. In such proposed systems,  
10 automatic speech recognition (ASR) systems are used in order to recognise voice inputs and produce the network location address on the basis of the recognition.

However, in order that a portable device with low  
15 processing power is able to implement such voice control, the vocabulary that is stored to be recognised will necessarily be small. Therefore it has been proposed that automatic speech recognition (ASR) be carried out by the network itself. However, this  
20 requires the opening of a voice channel between the portable device and the network. This has the obvious disadvantage that the voice channel itself may introduce distortion echoes or noise, thereby degrading the recognition performance. Also, the speed of the  
25 voice recognition will be severely compromised.

#### SUMMARY OF THE PRESENT INVENTION

Embodiments of the present invention can therefore  
30 provide a system for overcoming these disadvantages. In one particular embodiment, there is provided a method of communicating with a network having a plurality of applications, each of which has a location address on the network, the method comprising:  
35

defining a plurality of address components,  
respective combinations of which form location  
addresses;

storing reference voice signals;

5 associating the stored reference voice signals  
with respective address components;

receiving input voice signals;

comparing the input voice signals with the  
reference voice signals;

10 composing a location address from the address  
components that are associated with those reference  
voice signals which match the received input voice  
signals; and

communicating with the network on the basis of the  
15 composed location address.

According to another embodiment of the present  
invention, there is provided a method of controlling a  
service application in a mobile terminal, the service  
20 application being provided at a remote server having a  
location address in a network, wherein the method  
includes defining a plurality of address components  
which can be combined to form respective location  
addresses, and comparing input voice signals with  
25 stored reference voice signals associated with  
respective address components, thereby composing a  
location address for supply to the remote server from  
the mobile terminal.

30 It is emphasised that the term "comprises" or  
"comprising" is used in this specification to specify  
the presence of stated features, integers, steps or  
components, but does not preclude the addition of one  
or more further features, integers, steps or  
35 components, or groups thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic diagram illustrating access to a network;

5        Figure 2 illustrates parts of a device for accessing the network of Figure 1;

Figure 3 illustrates initial steps in a method embodying the present invention;

10       Figure 4 illustrates further steps in a method embodying the present invention; and

Figure 5 illustrates a device controlled by a server of the network of Figure 1.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

15       Referring to Figure 1, a network 1 is illustrated in which a web server 3 and a WAP server 4 are located. A particular example of the network 1 would be the internet, in which many such servers are provided for  
20       access by users. The connections within the network 1 are not shown in Figure 1 for the sake of clarity but it will be appreciated that communication to and from the servers 3 and 4 may be directly initiated from the user, or may be routed through various network links.  
25       Since the particular connection to the servers is not of concern to the present invention, the connections are not shown. As is known, the network 1 can be accessed by users. Two such users are illustrated in Figure 1. The first is a mobile user 5 which  
30       communicates with the network 1 via an RF communications link 6. The second user shown in Figure 1 is a fixed user 7 who communicates with the network over a fixed line access 8. Typically, the mobile user would access the network using a mobile telephone, a  
35       personal digital assistant, or laptop computer provided



with RF communications equipment. The fixed user 7 would typically use a personal computer (PC) to access the network 1.

5           As described above, recent standardization of the wireless application protocol (WAP) means that specific servers and locations can be provided for access by users.

10           As is known, each of the servers, and indeed each of the individual pages of information provided by the servers, have individual location addresses in the network. In order to access the information or  
15           application stored on a server, the user makes use of a browser application. The location address, for example URL, is supplied to the browser which accesses the network location indicated by the address. It will be appreciated that many of the location addresses can be complex, and not ideally suited to direct manual user  
20           input.

          Accordingly, in an embodiment of the present invention, the user terminal (such as the mobile user 5 or the fixed user 7) defines a plurality of address  
25           components. Respective combinations of the address components form individual location addresses for accessing the network. In order to access a location address, the user can then simply specify a combination of address components. In an embodiment of the present  
30           invention, reference voice signals (or voice marks) are stored by the user and are associated with respective address components. The reference voice signals can therefore be short and convenient labels for the address components. For example, the voice signal  
35           "web" could be associated with the address component

"http://www.", and the reference signal "home" could be associated with the user's preferred home page, for example "Ericsson.com/home". Thus, speaking "web, home" to the device would result, in this example, to production of the address "http://www.Ericsson.com/home", and the browser application would access that location on the network.

In this way, the user terminal need only compare input voice signals with the reference voice signals rather than performing full automatic speech recognition, in order to determine the location address.

Figure 2 illustrates the components of the user terminal for operation in accordance with the present invention. The terminal includes a processor 10, for carrying out the various data processing requirements of the device, and also includes a voice input unit 11, a data storage unit 12 and a communications unit 13. The voice input unit receives input voice signals from a user of the device and converts those signals into data for use by the processor 10. When the system is being set up or updated, the processor 10 stores the input voice signals and reference voice signals in the data storage unit 12. The processor then associates the reference voice signals with stored address components, for example storing the associations in a look-up table.

When the user wishes to access the network, he or she inputs a voice signal, via the voice input unit 11, to the processor 10. The processor 10 compares the input voice signal with the reference voice signals stored in the data storage unit 12, to determine which

of the reference voice signals match the input voice signals. The address components associated with the matched voice signals are retrieved from the storage unit 12, and are combined. The location address  
5 defined by the combination of address components is provided by the processor to the communications unit 13 for supply to the network 1 via a communications link 14. It will be appreciated that the communications link 14 may be an RF link in the case of a mobile user  
10 or a fixed link in the case of a fixed user.

Figure 3 is a flow diagram illustrating steps in an initial or updating routine in a method according to the present invention. At step A, the user (or the  
15 system) defines the address components that will be used in combination to compose location addresses. For example, the user may choose those components most suited to their use, or the system may select the most commonly used components. At step B, the user inputs  
20 and stores reference voice signals, or "voice marks", and at step C, the stored reference voice signals are associated with the address components defined in step A. The result of the method of Figure 3 is to produce a list of address components with their associated  
25 reference voice signals.

Operation of the system to access a location at a network will now be described with reference to Figure 4. The user terminal receives input voice signals  
30 (step D) and then compares those input signals with the stored reference voice signals (step E). The comparison of reference voice signals with the input voice signals will produce an indication of which reference voice signals match the input voice signals.  
35 From this indication, the address components indicated

by the input voice signals can be retrieved from the data storage unit 12 (Figure 2) for further processing (step F). When the address components have been retrieved, the location address to be accessed by the user can be composed (step G), and the network accessed (step H).

It will be readily appreciated that storing a plurality of address components with their associated references enables the overall amount of stored data to be reduced. This is because several of the individual components may be repeated in different addresses. For example the heading "http://www." is a very common component of locations on the internet. This component therefore needs only to be stored once for use in many different location addresses.

A particularly useful application of the present invention is for use in controlling devices, such as televisions, and other equipment, via a network location. For example, in Figure 5, the WAP server 4 of the network 1 of Figure 1 is shown connected via a controlled connection 17 to a control device 16. The control connection 17 can be provided by a fixed link such as a parallel port cable connection, or by a wireless connection such as the Bluetooth™ RF protocol. The example of Figure 5 shows a WAP server, but it is to be appreciated that the server can be any type of network server suitable for such an application. In order to control the control device 16, the user accesses the WAP server location using the browser in accordance with the present invention to provide commands for the WAP server to control the device 16. The particular location address supplied to the WAP server by the user terminal can issue the relevant

controls to the control device 16. Since many of the controls for the control device are likely to be common commands, each of these can be provided with an associated reference voice signal. All it is then  
5 necessary for a user to control a control device 16 via the server 4 is to speak voice control relating to the reference voice signals to the access device. This particular system is best described by reference to a particular example, as below.

10 For example, controlling household devices such as a television, a video and a lamp would require the user to define reference voice signals for the control of those devices. The voice signal "control" could  
15 provide the address component relating to the initial part of the WAP server address ("wapaddress:"), "TV" can define the address of the television on the WAP server, and "channel" to the computer "cgi?.channel=".. In addition, some generic components such as "ON"  
20 ("=on."), "OFF" ("=off.") and numbers 0 to 9 can be defined. Thus, to turn on the television to channel 3, the user need only speak "control TV ON channel 3", rather than "wapdevice://TVaddress/on.cgi?channel=3". It would be preferable for the user terminal to retain  
25 the last controlled device parameters (i.e. TV in the example) such that other controls can be shortcut further, so that changing the volume on the television would simply be a question of saying "volume up" or "volume down". These last voice inputs would then  
30 result in the addresses  
"wapdevice://TVaddress/cgi?volume=up" and  
"wapdevice://TVaddress/cgi?volume=down" respectively.

35 Building up the control language in the form of a location address for accessing a network in this way

from address components having their own reference voice signals, means that common commands need only be control 1. For example the numbers 0 to 9, on, off etc are likely to be common to the control of many devices. Also, it has the distinct advantage that the user need only speak short catchwords rather than needing to remember the full control address.

10 An idea complementing the present invention is to add voice tags to wireless mark-up language documents. Thus voice files can be attached to a location such that feedback (for example OK, error etc) can be supplied to the user. This is particularly suited to those occasions when a display is small or not provided.

20 It will be appreciated that the mobile users connection to the network can be via a mobile telephone network such as GSM or third generation GSM, or via a local RF communications system such as Bluetooth™. One advantage of the embodiment of the present invention is that standard communications protocols, such as Internet protocol, Wireless Application Protocol (WAP) or Bluetooth™ (standard communication protocol), can be used to control devices via a network server using standard access techniques. The voice activation of the access is enabled by the use of voicemarks. In the case of a mobile device accessing WAP pages, the device can provide the usual WAP page output to feedback to the user.

35 The present invention is particularly useful when a user is using a headset and does not have access to a keyboard. For example, a user may use a hands-free headset with a mobile telephone. In that case the user

does not have direct access to the keypad of the mobile telephone. The present invention then allows the user to access web pages and to control devices without reference to the keypad. Voice feedback from the page or device concerned removes the need for access to a display device.

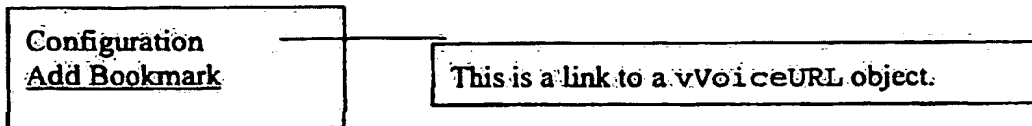
A further development of the invention enables complete complex URLs to be assigned voice bookmarks. One drawback of associating complex URLs with voice bookmarks is that the URL can be complex to input initially.

For example. A URL used to control a garage door may have CGI calls with complicated parameters, e.g. <http://myGarage/openPort.cgi?open=1&user=arne>.

Accordingly, a new object type "vVoiceURL" that is stored in the WAP server in the remote device is introduced. The format of the vVoiceURL object is:

```
BEGIN: vVoiceURL
URL:   http://myGarage/openPort.cgi?open=1&user=arne
END:   vVoiceURL
```

In a WAP page in the remote device (e.g. the garage WAP server) there is a link to this object:



When the user selects the link to the vVoiceURL object (i.e. the remote device action), the browser in

the user's device will download the object and see that it is a vVoiceURL object. The user's device will then start its voice recognition training and let the user speak a phrase into the phone. Once the user is

5 satisfied with the phrase, the bookmark is saved (with URL Http://myGarage/openPort.cgi?open=1&user=arne) and the voice tag is associated with the newly created bookmark.

10           Accordingly, the user is prompted to create a voice tag for a complex URL, without the need for creating the URL itself. The creation of the URL is performed by the server.



CLAIMS:

1. A method of communicating with a network having a plurality of locations, each of which has a location address on the network, the method comprising:  
5       defining an address component, which forms a location address;  
          storing a reference voice signal;  
          associating the stored reference voice signal with  
10       the address component;  
          receiving an input voice signal;  
          comparing the input voice signal with the reference voice signal; and  
          if the received voice signal matches the stored  
15       voice signal, communicating with the network on the basis of the location address.

2. A method of communicating with a network having a plurality of locations, each of which has a location address on the network, the method comprising:  
20       defining a plurality of address components, respective combinations of which form location addresses;  
          storing reference voice signals;  
25       associating stored reference voice signals with respective address components;  
          receiving input voice signals;  
          comparing the input voice signals with the reference voice signals;  
30       composing a location address from the address components that are associated with those reference voice signals which match the received input voice signals; and  
          communicating with the network on the basis of the  
35       composed location address.

3. A method of communicating with a network having a plurality of locations, each of which has a location address on the network, the method comprising:

receiving input voice signals;

5 comparing the input voice signals with stored reference voice signals, the stored reference voice signals having address components associated therewith;

10 composing a location address from address components that are associated with those reference voice signals which match the received input voice signals; and

communicating with the network on the basis of the composed location address.

15 4. A method as claimed in any one of the preceding claims, wherein communication with the network takes place over a mobile telephone network.

20 5. A method as claimed in any one of claims 1 to 4, wherein communication with the network takes place over a short range RF communications network.

25 6. A method as claimed in claim 5, wherein the short range RF communications network is a Bluetooth-enabled network.

30 7. A method as claimed in any one of the preceding claims, wherein communicating with a network serves to control a device.

35 8. A method of accessing a network using a network access application, the network having a plurality of locations which have respective location addresses in the network, the method including defining a plurality of address components, combinations of

which forming respective location addresses, assigning  
reference voice signals to the address components,  
comparing input voice signals with the reference voice  
signals and producing a network access address location.  
5 on the basis of address components associated with  
matched input and reference voice signals.

9. A method of associating voice signals to  
address components for use in a method as claimed in  
10 any one of the preceding claims, the method comprising;  
upon access to a location address by a user,  
prompting the user to input a voice signal;  
receiving a voice signal from the user;  
storing the received voice signal; and  
15 associating the stored voice signal with the  
accessed location address.

10. Apparatus for accessing locations on a  
network in which network locations have respective  
20 location addresses, the apparatus comprising:  
a voice input unit for receiving input voice  
signals to input voice data;  
a data storage unit;  
a processing unit operable to receive input voice  
25 data and to supply and retrieve data to and from the  
data storage unit; and  
a communications unit for communicating with a  
network,  
wherein the processing unit is operable to  
30 associate reference voice signals with location address  
components, and to compose a network access location  
address from components associated with reference voice  
signals that match input voice signals.

35 11. Apparatus as claimed in claim 10, wherein the

communications unit is operable to communicate with the network over a mobile telephone network.

5           12. Apparatus as claimed in claim 10, wherein the communication unit is operable to communicate with the network over a short range RF communications network.

10           13. Apparatus as claimed in claim 12, wherein the short range RF communications network is a Bluetooth-enabled network.

15           14. Apparatus as claimed in any one of claims 10 to 13, wherein the communications unit is operable to control a device.

            15. Apparatus as claimed in any one of claims 10 to 14, wherein the composed location address indicates a wireless application protocol server.

20           16. A mobile telephone including apparatus as claimed in any one of claims 10 to 15.

25           17. A WAP-enabled portable electronic device including apparatus as claimed in any one of claims 10 to 15.

            18. A portable electronic device including apparatus as claimed in any one of claims 10 to 15.

30           18. A device for accessing network information including apparatus as claimed in any one of claims 9 to 14.



Application No: GB 0018027.3  
Claims searched: 1-15

Examiner: Hannah Sylvester  
Date of search: 9 February 2001

## Patents Act 1977 Search Report under Section 17

### Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK CI (Ed.S): H4L (LRAB, LRAD, LRAX)

Int CI (Ed.7): G01L 5/06, G06F 3/16, 11/00, 13/00, 15/163, G10L 3/00, 3/02, 5/06, 9/00, 15/26, H04Q 3/00, 11/04, H04L 12/02, 12/18, 12/28, 12/46, 12/54, 12/56, 12/58, 12/64, 12/66, H04M 1/00, 1/26, 3/00, 3/42, 3/493, 11/00, 11/06, 15/00, 15/16.

Other: Online: WPI EPODOC JAPIO

### Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
X	EP1041779A2 IBM (see whole document)	At least: 1-3 and 8-10
X	JP10084425A NIPPON (see abstract)	At least: 1-3 and 8-10

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date, earlier than, the filing date of this application.